



Manuscript for review

**DEVELOPMENT OF A SELECTION SYSTEM FOR THE DETECTION OF L-RIBOSE ISOMERASE EXPRESSING MUTANTS OF *ESCHERICHIA COLI***

Journal:	<i>Applied Microbiology and Biotechnology</i>
Manuscript ID:	AMB-07-14756
Manuscript Category:	Original Paper
Date Submitted by the Author:	30-Mar-2007
Complete List of Authors:	De Muynck, Cassandra; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis Van der Borght, Jef; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis De Mey, Marjan; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis De Maeseneire, Sofie; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis Van Bogaert, Inge; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis Beauprez, Joeri; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis Soetaert, Wim; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis Vandamme, Erick; Ghent University, Laboratory of Industrial Microbiology and Biocatalysis
Keyword:	L-arabinose isomerase, L-ribose isomerase, directed evolution, screening, gene disruption

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

DEVELOPMENT OF A SELECTION SYSTEM FOR THE  
DETECTION OF L-RIBOSE ISOMERASE EXPRESSING MUTANTS  
OF *ESCHERICHIA COLI*

Cassandra De Muynck\*, Jef Van der Borght, Marjan De Mey, Sofie L. De Maeseneire,  
Inge N.A. Van Bogaert, Joeri Beauprez, Wim Soetaert and Erick Vandamme

*Laboratory of Industrial Microbiology and Biocatalysis, Department of Biochemical and  
Microbial Technology, Faculty of Bioscience Engineering, Ghent University, Coupure  
links 653, B-9000 Gent, Belgium*

\*Corresponding author. Tel.: +32-9-2646028; Fax.: +32-9-2646231.  
E-mail. Cassandra.DeMuynck@UGent.be

## Abstract

L-Arabinose isomerase (E.C. 5.3.1.14) catalyzes the reversible isomerization between L-arabinose and L-ribulose and is highly selective towards L-arabinose. By using a directed evolution approach, enzyme variants with altered substrate specificity were created and screened in this research. More specifically, the screening was directed towards the identification of isomerase mutants with L-ribose isomerizing activity. Random mutagenesis was performed on the *E. coli* L-arabinose isomerase gene (*araA*) by error-prone PCR to construct a mutant library. In order to enable screening of this library, a selection host was first constructed in which the mutant genes were transformed. In this selection host, the genes encoding for L-ribulokinase and L-ribulose-5-phosphate-4-epimerase were brought to constitutive expression and the gene encoding for the native L-arabinose isomerase was knocked-out. L-Ribulokinase and L-ribulose-5-phosphate-4-epimerase are necessary to ensure the channeling of the formed product, L-ribulose, to the pentose phosphate pathway. Hence, the mutant clones could be screened on a minimal medium with L-ribose as the sole carbon source. Through the screening, two first-generation mutants were isolated, which expressed a small amount of L-ribose isomerase activity.

# 1 Introduction

L-Arabinose isomerase (E.C. 5.3.1.4) is the first enzyme involved in the catabolism of L-arabinose by many microorganisms. The gene *araA*, encoding for this enzyme, is part of a larger gene complex which consists of 4 operons: *araBAD*, *araC*, *araE* and *araFG*. These genes are responsible for L-arabinose catabolism, gene regulation, low-affinity transport, and high-affinity transport, respectively (Lichenstein *et al.* 1987). The *araBAD* operon includes three structural genes encoding for L-arabinose isomerase, L-ribulokinase and L-ribulose-5-phosphate-4-epimerase, respectively. The end-product of this pathway, D-xylulose-5-phosphate, is subsequently taken up in the pentose phosphate cycle. The expression of these structural genes in *Escherichia coli* is regulated by AraC, the gene product of *araC*, by both positive and negative control mechanisms, involving the formation of alternating DNA loops (Huo *et al.* 1988). In the presence of L-arabinose, AraC protein ligated to L-arabinose, is required for transcription at the promoters of the *araBAD*, *araE* and *araFG* operons. In addition, AraC protein is also a repressor of *araBAD* and *araC* transcription, functioning to reduce the level of *araBAD* in uninduced cells and to autoregulate its intracellular level under both inducing and noninducing conditions (Lichenstein *et al.* 1987). Hence, the expression of the three structural genes involved in the metabolism of L-arabinose requires the presence of L-arabinose as inducer.

In our study, it is the final goal to evolve L-arabinose isomerase, by using directed evolution, to a variant isomerase, capable of isomerizing L-ribulose into L-ribose, hence an L-ribose isomerase.

Directed evolution is a technique, which uses the same algorithm used by nature to create an enormous biodiversity over millions of years, to evolve enzymes on laboratory scale in a much shorter time span, by introducing changes at DNA-level (mutagenesis) and identifying good mutants via high-throughput screening and selection (Bornscheuer 2005; Kaur *et al.* 2006). Hence, directed evolution comprises two distinct steps that are typically applied in an iterative fashion: (1) generating genetic diversity and (2) finding among the mutant sequences those that code for proteins that perform the desired function according to the specified criteria (Arnold & Georgiou 2003). In many ways, the second step is the most challenging. No matter how cleverly designed or diverse the starting library, without an efficient screening strategy the ability to isolate useful clones is severely diminished.

In our research, directed evolution was used to evolve L-arabinose isomerase to an L-ribose isomerizing variant, hence an L-ribose isomerase. For this purpose, the *araA* gene from *E. coli* MG1655 (= type strain) was cloned and over-expressed in *E. coli*. To avoid background activity disturbance, a knock-out mutant of *araA* was created in which the *araA* gene was over-expressed on a plasmid vector. This gene was diversified by using error-prone PCR to create a library of mutated isomerases, which was cloned in *E. coli*. Subsequently, this library was screened for the presence of L-ribose isomerase expressing clones by growing them in a minimal medium with L-ribose as the sole carbon source. For this purpose, a selection host (*E. coli*) was first constructed in which the genes for L-ribulokinase (*araB*) and L-ribulose-5-phosphate-4-epimerase (*araD*) were constitutively expressed by changing the natural, L-arabinose inducible  $P_{BAD}$  promoter with a constitutive one ( $P_{37}$ ). Hence, L-ribulose, formed by the action of the putative L-ribose

isomerase, could be further metabolized without the need for addition of L-arabinose as inducer, as this would undo the effect of the minimal screening medium with L-ribose. With this selection host and minimal medium, it was possible to identify L-ribose isomerase expressing clones from the library of mutated L-arabinose isomerases. The improvement of their activity through iterative cycles of directed evolution, should lead to an enzyme that is capable of producing L-ribose from L-ribulose. L-Ribose is an important lead molecule in the synthesis of antiviral nucleoside-analogue drugs (Ahmed *et al.* 1999).

## 2 Materials and Methods

### 2.1 Chemicals and enzymes

L-Arabinose, bovine serum albumin (BSA), Bradford reagent, carbenicillin, carbazol, chloramphenicol, cystein.HCl, IPTG kanamycin, L-lactate dehydrogenase, pyruvate kinase, phosphoenolpyruvate, NADH and tetracycline were purchased from Sigma-Aldrich (Belgium). L-Ribulose was synthesized by dehydrogenation of ribitol (De Muynck *et al.* 2006). L-Ribose was a kind gift of CMS Chemicals (UK). Yeast extract, tryptone and agar were obtained from Difco (Belgium). Restriction enzymes and *Taq* polymerase were obtained from New England Biolabs - Westburg (The Netherlands). In-Fusion™ PCR Cloning Kit was purchased from Clontech-Takara BIO (France) and GeneMorph® II EZClone Domain Mutagenesis Kit was purchased from Stratagene (The Netherlands), High Fidelity PCR Master Kit was purchased from Roche Applied Science (Belgium). Sequence alignments were made and PCR primers were designed with Clone Manager Professional 8 (Sci Ed Software, USA) and synthesized by Sigma-Genosys

(Belgium). Routine recombinant DNA methodology was performed according to Sambrook and Russell (2001). Nucleotide sequencing was performed by the VIB Genetic Service Facility (Belgium).

## 2.2 Strains, plasmids and growth conditions

*E. coli* strain MG1655 was used as wild-type strain for the creation of the selection host, named LRISel2. The expression vector pOExA contains an ampicillin resistance gene, a multi-cloning site, a *lac* repression system and the strong *trc* promoter. *E. coli* strains were grown in Luria-Bertani Broth (LB; 10 g L<sup>-1</sup> trypton, 5 g L<sup>-1</sup> yeast extract and 10 g L<sup>-1</sup> NaCl) at 37 °C with vigorous shaking.

## 2.3 Construction of the selection host *E. coli* LRISel2

For the construction of the selection host, *araA* was knocked-out and *P<sub>BAD</sub>* was substituted by the method for gene disruption of Datsenko and Wanner (Datsenko & Wanner 2000). The basic strategy is to replace a chromosomal sequence with a selectable antibiotic resistance gene and the insertion fragment that is generated by PCR by using primers with circa 36 nt homology extension (H1 and H2). After selection, the resistance gene can also be eliminated by using a helper plasmid expressing FLP recombinase, which acts on the directly repeated FRT-sites flanking the resistance gene. The Red and FLP helper plasmids can easily be cured by growth at 42 °C because they have temperature-sensitive replicons. The plasmids used in this strategy, pKD46, pKD3, pKD4, and pCP20, were a kind gift from prof. R. Cunin (Laboratory for Microbiology and Genetics, Free University of Brussels, Belgium), whereas p37-pKD3 and p37-pKD4 were constructed by De Mey et al. (2007), based on the former plasmids. pKD46 (Amp<sup>R</sup>)

and pCP20 (Amp<sup>R</sup>) are easily curable, low copy number plasmids, carrying the genes for λ Red recombinase and FLP recombinase, respectively. pKD3 (Cm<sup>R</sup>) and pKD4 (Km<sup>R</sup>) contain an FRT-flanked antibiotic resistance gene for chloramphenicol (*cat*) and kanamycin (*kan*) resistance.

2.4 Cloning of the full-length *araA* gene

Based on the nucleotide sequence stored in the KEGG database (eco:b0062, Blattner et al., 1997), the genomic DNA encoding *E. coli* L-arabinose isomerase was isolated by PCR with two synthetic primer pairs (*pr\_araA\_1*: 5'- GCCAGGTTGGCTTCTAATAC-3'; *pr\_araA\_2*: 5'- CGAAAGTGCACGCAGACATC-3'; *infusion\_araA\_1*: 5'- AGGAAACAGACCATGACGATTTTTGACAACCTATGAAGTGTG-3'; *infusion\_araA\_2*: 5' – AGGATCCCCCCCCATGTTAGCGACGAAACCC-3'), the first pair serving to pick-up the gene from the *E. coli* MG1655 genome and the second pair to create the necessary homology for the insertion in pOExA. The PCR product was ligated with the In-Fusion PCR Cloning Kit (Clontech-Takara BIO, France) into the pOExA expression vector, previously digested with *Nco*I, and the ligation mixture was transformed into *E. coli* FusionBlue and LRISel2. The resulting recombinant plasmid, named pAIIF, was confirmed by restriction mapping, PCR (with *control ara forw2* 5'- ATGGCTGTGCAGGTCGTAAATC -3'; *control ara rev2* 5'- ACGGCGTTTCACTTCTGAGTTC-3') and nucleotide sequencing.

2.5 Random mutagenesis

Random mutagenesis of the *araA* gene was performed by error-prone PCR with the GeneMorph<sup>®</sup> II EZClone Domain Mutagenesis Kit (Stratagene). The error prone PCR



was performed using Mutazyme II DNA polymerase with two synthetic oligomers (*devfwd3*: 5'-CACAGGAAACAGACCATG-3'; *devrev3*: 5'-TCTAGAGGATCCCCGCCATGTTA-3') and the amplified error-prone PCR product was used as megaprimer for the EZClone reaction. The resulting amplified plasmid was digested by *DpnI* and transformed into *E. coli* LRISel2 by electroporation. The transformants were incubated for 1 hour in LB broth at 37 °C and subsequently thoroughly washed with ice-cold distilled water before plating them on the selection medium.

## 2.6 Selection of L-ribose isomerase expressing clones

The selection medium contained  $\text{NH}_4\text{Cl}$  (2 g  $\text{L}^{-1}$ ),  $(\text{NH}_4)_2\text{SO}_4$  (5 g  $\text{L}^{-1}$ ),  $\text{KH}_2\text{PO}_4$  (3 g  $\text{L}^{-1}$ ),  $\text{K}_2\text{HPO}_4$  (7.3 g  $\text{L}^{-1}$ ),  $\text{NaCl}$  (0.5 g  $\text{L}^{-1}$ ),  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  (0.5 g  $\text{L}^{-1}$ ), L-ribose (16.5 g  $\text{L}^{-1}$ ) and 1 mL  $\text{L}^{-1}$  of a 1000x vitamin-mineral solution. The vitamin-mineral solution consisted of  $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$  (3.6 mg  $\text{L}^{-1}$ ),  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  (5 mg  $\text{L}^{-1}$ ),  $\text{MnCl}_2 \cdot 2\text{H}_2\text{O}$  (1.3 mg  $\text{L}^{-1}$ ),  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$  (0.38 mg  $\text{L}^{-1}$ ),  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$  (0.5 mg  $\text{L}^{-1}$ ),  $\text{ZnCl}_2$  (0.94 mg  $\text{L}^{-1}$ ),  $\text{H}_3\text{BO}_4$  (0.03 mg  $\text{L}^{-1}$ ),  $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$  (0.4 mg  $\text{L}^{-1}$ ),  $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$  (0.0967 mg  $\text{L}^{-1}$ ) and thiamine.HCl (1.01 mg  $\text{L}^{-1}$ ). Induction was performed by adding 1 mL  $\text{L}^{-1}$  of a 1000x IPTG stock solution (100 mM).

## 2.7 Enzyme assays

### 2.7.1 Cell harvest and disruption for enzyme assays

*E. coli* cells were harvested 6 hours after IPTG induction for isomerase assays or after overnight culturing for ribulokinase assays, by centrifugation and the cellular pellets were

frozen until further use. Frozen pellets were then resuspended in 25 mM Tris buffer (pH 7.0) supplemented with EasyLyse™ Bacterial Protein Extraction solution (Epicentre – BIOzymTC, The Netherlands) for lysis. Supernatant was obtained after centrifugation and used as crude cell extract for both enzyme assays and protein determination.

**2.7.2 Isomerase assay**

The activity of L-arabinose isomerase was assayed as described by Roh and coworkers (2000). The reaction progress was monitored by measuring the L-ribulose content with the spectrophotometric method of Dische and Borenfreund (1951). The concentration was calculated following a previously determined standard curve:  $OD_{540} = 1.01 \cdot 10^{-2} \times (\text{mg L}^{-1}) + 1.31 \cdot 10^{-2} R^2 = 0.9996$ . One unit of enzyme is expressed as the amount of enzyme capable of producing 1 µg of L-ribulose per minute. Protein content was determined according to the procedure of Bradford (1976), using BSA as a standard:  $OD_{595} = 3.72 \cdot 10^{-3} \times (\text{mg dL}^{-1}) + 1.91 \cdot 10^{-1} R^2 = 0.9959$ .

**2.7.3 L-Ribulokinase assay**

For the measurement of L-ribulokinase activity, a spectrophotometric method adapted from Lee and Englesberg (1962), was used. Practically, 30 µL of crude cell extract were added to 155 µL of a solution containing 21.2 mM potassium phosphate, 0.34 mM phosphoenolpyruvate, 0.34 mM L-ribulose, 0.16 mM β-nicotinamide adenine nucleotide (reduced form), 5.3 mM magnesium sulphate, 1.06 mM adenosine-5'-diphosphate, 5.3 units L-lactic dehydrogenase and 0.01 units of pyruvate kinase. L-ribulokinase activity was determined by measuring the disappearance of NADH at 340 nm continuously during 5 min. at 37 °C.

### 3 Results and discussion

#### 3.1 Creation of a selection host for directed evolution of L-arabinose isomerase

In this study, it was the aim to evolve L-arabinose isomerase from *E. coli* to an L-ribose isomerizing mutant, which was also to be expressed in *E. coli*. Therefore, mutants were selected on a minimal medium containing L-ribose as the sole carbon source. As such, only the good mutants survive and a massive number of bad clones is immediately excluded of further assays. In wild-type *E. coli*, L-ribulose (generated from L-arabinose by wild-type L-arabinose isomerase) is further metabolized by phosphorylation with L-ribulokinase (AraB) and epimerization with L-ribulose-5-phosphate-4-epimerase (AraD), as shown in Figure 1. Hence, if L-ribulose (originated from L-ribose) needs to be converted to D-xylulose-5-phosphate, *araB* and *araD* should be expressed. The genes encoding for these enzymes are located in the same transcription unit as *araA*, under the control of the L-arabinose inducible  $P_{BAD}$  promoter. Therefore, expression of *araB* and *araD* can be achieved by addition of L-arabinose to the growth medium, which activates the regulator AraC. Since in the selection system, L-arabinose can not be used as an inducer for the transcription of these genes, the  $P_{BAD}$  encoding region was knocked out and substituted with the strong constitutive  $P_{37}$  promotor, constructed by De Mey et al. (2007).

In order to exclude background activity of *E. coli*'s wild-type genomic L-arabinose isomerase the *araA* gene was also knocked-out. This created a selection strain that is not longer capable of catabolizing L-arabinose when the expression vector, containing L-

arabinose isomerase, is absent and gave the opportunity to detect those mutations that do not influence the activity of L-arabinose isomerase after error-prone PCR and those that do change the substrate specificity. An overall scheme of the constructed selection host is given in Figure 2.

Knock-out of *araA* and knock-in of  $P_{37}$  were achieved with the method of Datsenko and Wanner (2000) by creating PCR products that have DNA ends homologous to the region to disrupt. Based on the nucleotide sequence and organization of *araBADaraC* obtained from EcoCyc (EcoCyc.org), appropriate homology regions were identified. It is important to note that the transcription start site of  $P_{BAD}$  had to be included in the homology region for the substitution with  $P_{37}$ , since the natural ribosome binding site is more effective and stable to promote translation (De Mey *et al.* 2007). The selected homology regions used to design the primers are shown in Figure 3. These homology regions were used to create primers for amplification of the desired plasmid DNA (pKD3/pKD4 and p37-pKD3/pKD4). Therefore, oligonucleotides that prime with these plasmids were combined with the homology regions as overhang. Hence, PCR with these combined primers (as shown in Table 1) creates a linear DNA fragment that can be recombined in the genome of *E. coli* to disrupt *araA* and  $P_{BAD}$ .

**3.1.1 Knock out *AraA***

PCR of pKD3 and pKD4 with primers KnockOut\_AraA\_H1P1 and H2P2 created a 1086 bp and 1549 bp product, respectively. These products were both transformed separately into an L-arabinose induced, pKD46 possessing *E. coli* MG1655 and incubated for the recombination to proceed. After spreading the transformants on a selective agar plate, several colonies were tested with colony PCR, confirming the insertion of the FRT-

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

flanked resistance gene. The mutants were colony-purified, incubated at 42 °C and tested for loss of the pKD46 plasmid. After transformation with pCP20 and incubation to cut-out the antibiotic resistance gene, transformants were again spread on selective agar and verified by using the same strategy. Finally, good mutants containing the *cat* gene were isolated and the absence of *araA* in these mutants was confirmed by PCR with the control\_KOaraA primers and by sequencing of the obtained PCR product.

### 3.1.2 Knock in $P_{37}$

The same strategy was followed to replace  $P_{BAD}$  with  $P_{37}$ . PCR of p37-pKD3 and p37-pKD4 with primers KnockIn\_P37\_H1P1 and H2P2 yielded a 1620 bp and 1157 bp product, respectively. These products (which consist of the  $P_{37}$  sequence followed by an FRT-flanked resistance gene) were transformed separately into an L-arabinose induced, pKD46 possessing *E. coli* ( $\Delta araA$ ) and incubated for the recombination to proceed. After spreading the transformants on a selective agar plate, several colonies were tested with colony PCR, which yielded an 1157 bp long fragment for the *cat*-containing clones, confirming the insertion of the FRT-flanked resistance gene. The mutants were colony-purified, incubated at 42 °C and tested for loss of the pKD46 plasmid. The success of the promoter substitution was confirmed by PCR with the control primers and by nucleotide sequencing. It is noteworthy that the FRT-flanked resistance gene was not yet eliminated at this stage. Finally, one of these mutants was selected as selection host and named LRISel2.

3.1.3 Enzyme assays

An assay for the measurement of L-arabinose isomerase activity in LRISel2 showed almost no production of L-ribulose from L-arabinose ( $0.11 \pm 0.03 \text{ U mg}^{-1}$ ), whereas in the wild-type an activity of  $4.20 \pm 1.05 \text{ U mg}^{-1}$  was noticed. This confirms too the successful knock out of *araA*.

An enzyme assay for the measurement of L-ribulokinase was also performed. This is the first enzyme after L-arabinose isomerase involved in L-arabinose catabolism, and was brought to constitutive (over)expression by insertion of the *P*<sub>37</sub> promoter in the place of the natural, L-arabinose inducible *P*<sub>BAD</sub> promoter. With this assay (Fig. 4), it could be determined that the LRISel2 strain showed a 2-fold increase in L-ribulokinase activity as compared to the wild-type MG1655 strain. However, it is notheworthy that, even without induction of L-arabinose, a low level of L-ribulokinase is also present in MG1655.

3.2 Cloning and expression of the full-length *araA* gene

In a first step, *araA* was amplified from genomic DNA of *E. coli* MG1655 with primers pr\_araA\_1 and 2. The success of the PCR was confirmed by electrophoresis, which showed a strong signal at 2041 bp. In the next step, the gene was cloned into the pOExA expression vector with the aid of the In-Fusion PCR Cloning Kit. For this purpose, the PCR product was amplified again, using the infusion\_araA\_1 and infusion\_araA\_2 primers, which contained both a homology region to the *araA* gene and to the vector site (*NcoI*) where it had to be inserted. This PCR product was then ligated in pOExA and named pAIIF (L-Arabinose Isomerase InFusion). This plasmid was transformed in *E. coli* FusionBlue competent cells provided with the kit and subsequently in the selection host

*E. coli* LRISel2. The clone expressing the strongest signal in PCR - gel electrophoresis was confirmed by restriction mapping and nucleotide sequencing of its plasmid and was used in an L-arabinose isomerase activity assay. The thus measured activity was  $102.36 \pm 11.73 \text{ U mg}^{-1}$ , which confirmed the success of the cloning and expression of *araA*. As a comparison, the wild-type strain MG1655 and the  $\Delta araA$  strain LRISel2 without expression vector showed almost no activity.

Finally, the constructed selection strain was tested for growth on the selection medium, which contained only L-arabinose as C-source. The wild-type strain *E. coli* MG1655 grew very well on the selection medium, whereas the constructed LRISel2 strain without *araA*, did not. However, upon complementing this strain with the constructed pAIF 4-2, which contains the *araA* gene, the capability to grow on this medium, was fully restored. This proves again that the knock-out was successfully constructed.

### 3.3 Mutagenesis

Error-prone PCR was used to create a library of mutated *araA* genes, from which L-ribose isomerizing clones were isolated. The library was made by using the GeneMorph® II EZClone Domain Mutagenesis Kit (Stratagene), and the (theoretical) mutation frequencies used were 2, 6 and 10 mutations per kb. After transformation, the clones were used to inoculate three liquid selection medium cultures and after 72 hours of incubation, OD<sub>660</sub> had increased to almost 2.0 in all three cultures. The resulting activity, measured both with L-ribose (new substrate) and L-arabinose (natural substrate) is shown in Figure 5.

As depicted in the graph, L-ribose isomerizing activity could be detected in the 6 and 10 mutations per kb samples. At the same time, these samples lost their ability to isomerize L-arabinose, the natural substrate of the wild-type L-arabinose isomerase. For the 10 mutations per kb sample, the L-ribose isomerase activity reached 13 units per mg protein. Practically, this means that 0.17 mg of L-ribulose were formed after 90 minutes of incubation from 20 mg of L-ribose, or 0.84 % conversion compared to 18.1 % conversion of L-arabinose by the wild-type pAIF 4-2 containing strain and  $\approx 0.01$  % conversion of L-ribose by this strain. In directed evolution experiments, it is not unusual that the first generation of mutants expresses only a very small amount of the “new” activity. This is one of the reasons why the directed evolution approach is often performed in an iterative manner, where the activity is increased in subsequent mutation rounds (Arnold & Georgiou 2003).

It is noteworthy that in the experiment, the cell extract of the mutant cultures contained only half as much protein than the one of the wild-type culture, hence the latter contained more biocatalyst. The equilibrium ratio of the reaction is 70:30 (28 °C, pH 9.0) for L-ribose:L-ribulose (Ahmed *et al.* 1999) and varies between 90:10 (25 °C, pH 7.0) and 84:16 (53 °C, pH 7.0) for L-arabinose:L-ribulose, as determined by Heath et al (1958) with L-arabinose isomerase from *Lactobacillus plantarum*. This indicates that, with the wild-type enzyme the chemical equilibrium is already reached almost after 90 minutes of incubation, whereas in the selected mutant culture it is not. Hence, there is still much room for improvement and further mutagenesis of the obtained enzyme is needed. However, the fact that there is some L-ribose isomerase activity present, whereas in the



wild-type there is none, indicates that one or more amino acid mutations were introduced that impact on the substrate specificity of the isomerase.

## 4 Conclusion

L-Arabinose isomerase (AraA) is the first enzyme involved in the catabolism of L-arabinose by many microorganisms. It converts L-arabinose to its keto-derivative, L-ribulose, which is in turn phosphorylated by L-ribulokinase (AraB) and epimerized to D-xylulose-5-phosphate by L-ribulose-5-phosphate-4-epimerase (AraD).

In this research, directed evolution was used to evolve L-arabinose isomerase to an L-ribose isomerizing variant, hence an L-ribose isomerase. For this purpose, the *araA* gene from *E. coli* MG1655 (= type strain) was cloned and over-expressed, again in *E. coli*. To avoid background activity disturbance, a knock-out mutant of *araA* was created in which the *araA* gene was over-expressed on a plasmid vector. This gene was diversified by using error-prone PCR to create a library of mutated isomerases, which was cloned in *E. coli*. Subsequently, this library was screened for the presence of L-ribose isomerase expressing clones by growing them in a minimal medium with L-ribose as the sole carbon source. For this purpose, a selection host (*E. coli*) was first created in which the genes for L-ribulokinase (*araB*) and L-ribulose-5-phosphate-4-epimerase (*araD*) were brought to constitutive expression by changing the natural, L-arabinose inducible  $P_{BAD}$  promoter for a constitutive one ( $P_{37}$ ). Hence, L-ribulose, created by the action of the putative L-ribose isomerase, could be further metabolized without the need for addition of L-arabinose as inducer, as this would undo the effect of the minimal screening medium with L-ribose.

With this selection host and minimal medium, it was possible to propagate L-ribose isomerase expressing clones from the library of mutated L-arabinose isomerases.

## 5 Acknowledgements

The authors would like to thank the Special Research Fund (BOF) of Ghent University for their financial support.

## 6 References

Ahmed, Z., Shimonishi, T., Bhuiyan, S. H., Utamura, M., Takada, G., Izumori, K. (1999). Biochemical preparation of L-ribose and L-arabinose from ribitol: A new approach. *Journal of Bioscience and Bioengineering*, 88, 444-448.

Arnold, F. H., Georgiou, G. (2003). Directed enzyme evolution: Screening and selection methods. Totowa, New Jersey, Humana Press Inc., 383 p.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Biochemistry*, 72, 248-254.

Datsenko, K. A., Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of The National Academy of Sciences of The United States Of America*, 97, 6640-6645.

De Mey, M. (2007). Metabolic modeling and engineering of *Escherichia coli* to minimize acetate formation in recDNA fermentations. PhD thesis, Ghent University, 220 p.

De Mey, M., Maertens, J., Lequeux, G., Soetaert, W., Vandamme, E. (2007). Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. *BMC Biotechnology*, submitted.

De Muynck, C., Pereira, C., Soetaert, W., Vandamme, E. (2006). Dehydrogenation of ribitol with *Gluconobacter oxydans*: Production and stability of L-ribulose. *Journal of Biotechnology*, 125, 408-415.

Dische, Z., Borenfreund, E. (1951). A new spectroscopic method for the detection and determination of keto sugars and trioses. *Journal of Biological Chemistry*, 192, 583-587.

Heath, E. C., Horecker, B. L., Smyrniotis, P. Z., Takagi, Y. (1958). Pentose fermentation by *Lactobacillus plantarum*, II. L-Arabinose isomerase. *Journal of Biological Chemistry*, 231, 1031-1037.

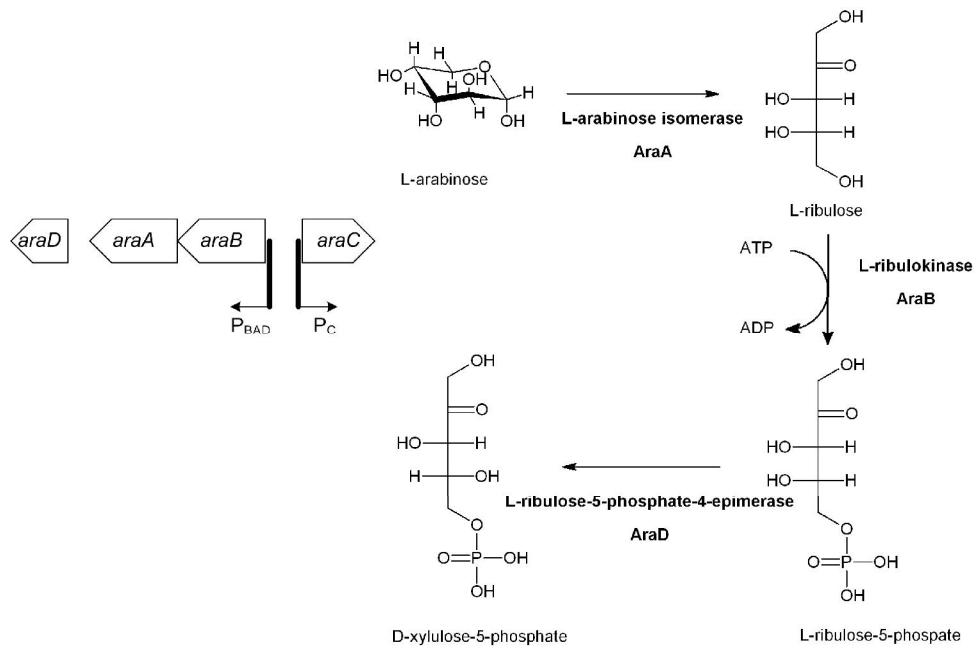
Huo, L., Martin, K. J., Schleif, R. (1988). Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. *Proceedings of The National Academy of Sciences of The United States Of America*, 85, 5444-5448.

Lee, N., Englesberg, E. (1962). Dual effects of structural genes in *E. coli*. *Journal of Bacteriology*, 48, 335-348.

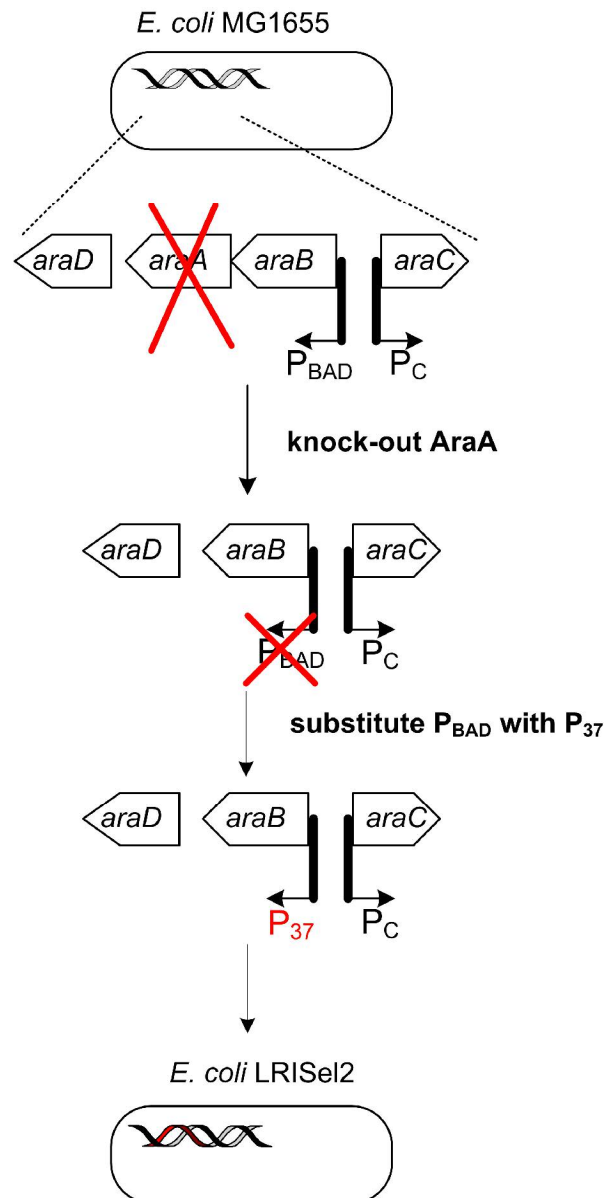
Lichenstein, H. S., Hamilton, E. P., Lee, N. (1987). Repression and catabolite gene activation in the *araBAD* operon. *Journal of Bacteriology*, 169, 811-822.

Roh, H.-J., Yoon, S.-H., Kim, P. (2000). Preparation of L-arabinose isomerase originated from *Escherichia coli* as a biocatalyst for D-tagatose production. *Biotechnology Letters*, 22, 197-199.

Sambrook, J., Russell, D. (2001). Molecular cloning: a laboratory manual. New York, Cold Spring Harbor Laboratory Press, p.

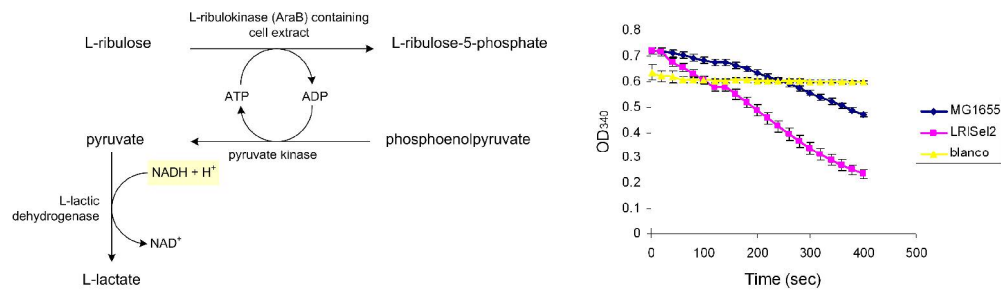


**Figure 1: The first three genes in L-arabinose metabolism are associated in an operon under control of an AraC inducible promoter; the regulator AraC participates also in controlling genes involved in L-arabinose transport**

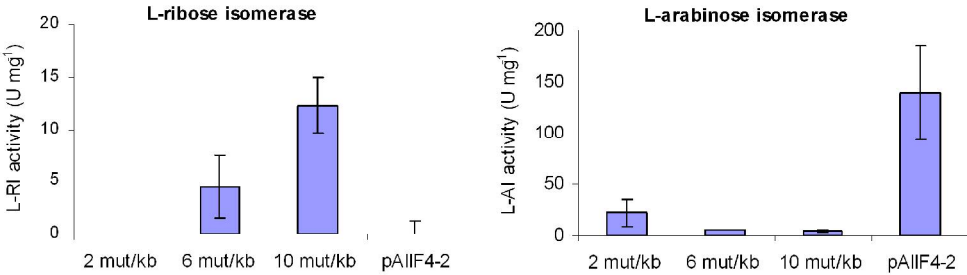


**Figure 2: Scheme of the constructed selection host, *E. coli* LRISel2**

**Figure 3: Organization of araBADaraC operon in *E. coli* (source: EcoCyc.org) with indication of the chosen homology regions for the disruption of *araA* and *P<sub>BAD</sub>***



**Figure 4: L-Ribulokinase assay: assay design and result with cell extracts of *E. coli* MG1655 and LRISel2, compared to a blank which contained buffer instead of cell extract**



**Figure 5: Isomerase activity of mutants selected in liquid minimal medium with L-ribose as the sole carbon source; 1 unit of enzyme is expressed as the amount of enzyme capable of producing 1 µgram of L-ribulose per minute; L-RI = L-ribose isomerase, L-AI = L-arabinose isomerase**



**Table 1: Sequence of primer pairs used in this study; homology regions are indicated underlined**

Aim	Name	Sequence (5'-3')
Knock out <i>araA</i>	KnockOut_AraA_H1P1	<u>TCAGCGTCGCATCAGGCGTTACATACCGGATGCGGCGTG</u> TAGGCTGG AGCTGCTTC
	KnockOut_AraA_H2P2	<u>GGCACAGGCTGCCCAGGCCGTTGCGACTCTATAAGGCATATGAATAT</u> CCTCCTTAG
Verify knock out	Control_KOaraA_fwd5	GCCAGGTTGGCTTCTAATAC
	Control_KOaraA_rev5	CGAAAGTGCACGCAGACATC
Knock in <i>P37</i>	KnockIn_P37_H1P1	<u>GCAATCGCCATCGTTTCACTCCATCCAAAAAACGGGTAAAAATGAC</u> ATATACCACATGGA
	KnockIn_P37_H2P2	<u>CCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCCAT</u> ATGAATATCCTCCTTAG
Verify knock in	Control_KIP37_fwd	TGCAATCGCCATCGTTTCAC
	Control_KIP37_rev	TAACAAAGCGGGACCAAAGC